THE FAILURE RATE AND RESIDUAL MIXING PROBABILITIES RESULTING FROM A MIXTURE OF DISTRIBUTIONS

Michael Wayne O'Bar

NAVAL POSTGRADUATE SCHOOL

Monterey, California

THESIS

THE FAILURE RATE AND
RESIDUAL MIXING PROBABILITIES
RESULTING FROM A MIXTURE OF DISTRIBUTIONS

by

Michael Wayne O'Bar

Thesis Advisor: J. D. Esary

September 1972

The Failure Rate and Residual Mixing Probabilities Resulting From a Mixture of Distributions

by

Michael Wayne O'Bar
Lieutenant, United States Navy
B.S., University of San Francisco, 1965

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 1972

ABSTRACT

Properties are developed for the failure rate function and residual mixing probabilities resulting from a mixture of distributions. After a detailed investigation of these properties for purely exponential mixtures, two more general mixtures are considered. Emphasis is placed on relating the current failure rate in the mixed population to the population's residual composition. The results have potential application in determining an appropriate period of "burn in."

TABLE OF CONTENTS

I.	INTRODUCTION	5
II.	THE CASE OF TWO EXPONENTIALS (DISCRETE PARAMETERS)	8
III.	MIXTURE OF THREE EXPONENTIAL TYPES (DISCRETE PARAMETERS)	18
IV.	GENERAL MIXTURE OF EXPONENTIAL TYPES (DISCRETE PARAMETERS)	23
V.	CONTINUOUS PARAMETER EXPONENTIAL MIXTURES	28
VI.	SOME NONEXPONENTIAL MIXTURES	32
BIBLIOGRAPHY		40
INITIAL DISTRIBUTION LIST		41
FORM DD 1473		43

LIST OF DRAWINGS

1. Two Exponential Failure Rate Functions -------------- 9

2. Failure Rate Function Resulting from a Mixture of Two Exponential Types ($p_1 < p_2$)-----------13

I. INTRODUCTION

Suppose that at time t=0 there exists a population of similar elements. This population could be made up of a type of transistor, a type of hydraulic check valve, a certain type of lock, *et cetera*. But whatever the elements composing the population may be, they are all designed to perform the same task. And being a functional physical object, each element will have some random lifetime, depending on a known or unknown probability distribution, during which it functions properly. When this random lifetime has elapsed, the element fails.

The usual definition of the probabilistic failure rate function of an element with random lifetime will be used throughout this thesis. Namely, the failure rate function of an element is the density function for the lifetime of the element divided by the element's survival function. The scenario can now be further developed by requiring that the population consist of subgroups of elements, with each subgroup having its own unique failure rate function applicable to each element in the subgroup. There is, then, a mixture of failure rates in the population. It will be assumed that at time t = 0 the exact proportion of the total population held by each subgroup is known either explicitly in the discrete case or implicitly through a cumulative distribution function in the continuous case. And so, the proportion of the population at time t = 0 having a particular failure rate is known.

Now that the population has been defined, it is possible to ask the two questions whose answers this thesis is concerned with. First, if an element is selected at random from the population at time $t = 0$, what is the nature of the randomly selected element's failure rate function? And given that a particular element selected at random from the population at time $t = 0$ is still functioning at time $t > 0$, what is the probability that that element was a member of a specific subgroup?

For someone wanting an explanation of the practical importance of these questions, a plausible example might be the following. Suppose that the population consists of transistors which were not marked as to subgroup membership due to some oversight during the production process. If the installation of one of the transistors into a piece of electronic equipment at time $t = 0$ is required, what can be predicted about the failure rate of the installed transistor? And on the basis of its performance, what can be said about that transistor's subgroup source?

Others have already developed some of the answer to the first question raised above about the nature of the failure rate of an element selected at random from the population at time $t = 0$. For instance, Barlow, Marshall, and Proschan [Ref. 1] have shown that a mixture of distributions each having a nonincreasing failure rate itself has a nonincreasing failure rate. Proschan [Ref. 2] applies this result to help explain an observed phenomenon in the aircraft industry and

to conclude that a mixture of exponentials will have a decreasing failure rate. Esary, Marshall, and Proschan [Ref. 3] show that a mixture of distributions with decreasing hazard rate average has a decreasing hazard rate average. And Gnedenko et al. [Ref. 4] indicate that in a mixture of exponentials the failure rate always approaches as a limit, with increasing time, the least exponential parameter value of all elements composing the population. This thesis sheds further light on the nature of the failure rate of an element drawn at random from a population composed of elements with a mixture of failure rates. No claim is made that the exposition here begins to give a complete answer, and areas needing further study are indicated.

The second question, about the conditional probability that a randomly selected element has come from a subgroup with a particular failure rate function given that the element is still functioning at a certain time, does not seem to have been given much attention in the literature. And once again, the results presented here are certainly not exhaustive. More research is required if the question is to be completely answered.

Briefly, the procedure followed in this thesis is to show results for increasingly more complex mixtures of exponentials. Once this has been done, some attention is given to more general mixtures.

II. THE CASE OF TWO EXPONENTIALS (DISCRETE PARAMETERS)

Let the population at time $t = 0$ consist of a proportion p_1 of elements distributed exponentially with parameter λ_1 and a proportion p_2 of elements distributed exponentially with parameter λ_2. The equality $p_1 + p_2 = 1$ holds. Assume that λ_1 is less than λ_2. It is a well known result that the failure rate functions $h_1(t)$ and $h_2(t)$ for elements in the two respective population subgroups are given by $h_1(t) = \lambda_1$ and $h_2(t) = \lambda_2$ (see Figure 1).

At time $t = 0$ an element is selected at random from the population. The following definitions are made:

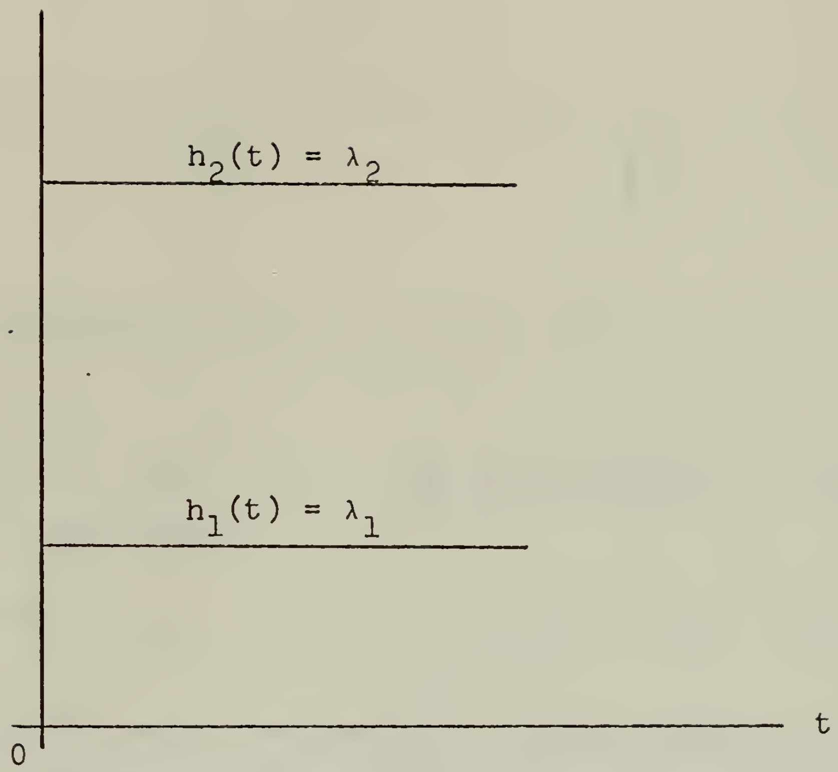
> A_1 is the event that the element selected comes from the subgroup with failure rate λ_1. A_2 is the event that the element selected comes from the subgroup with failure rate λ_2. T is the lifetime of the element selected at random from the population. And $\bar{F}_T(t)$, $f_T(t)$, and $h_T(t)$ are the element's survival, density, and failure rate functions, respectively.

Using these definitions, some equations can be derived:

$$\bar{F}_T(t) = P(T > t) = P(T > t|A_1)P(A_1) + P(T > t|A_2)P(A_2)$$

$$= p_1 e^{-\lambda_1 t} + p_2 e^{-\lambda_2 t} .$$

$$f_T(t) = -\bar{F}_T{}'(t) = p_1\lambda_1 e^{-\lambda_1 t} + p_2\lambda_2 e^{-\lambda_2 t} .$$



Two Exponential Failure Rate Functions

Figure 1

$$h_T(t) = \frac{f_T(t)}{\bar{F}_T(t)} = \frac{p_1\lambda_1 + p_2\lambda_2 e^{(\lambda_1 - \lambda_2)t}}{p_1 + p_2\, e^{(\lambda_1 - \lambda_2)t}} .$$

$$h_T'(t) = \frac{-p_1 p_2(\lambda_1 - \lambda_2)^2\, e^{(\lambda_1 - \lambda_2)t}}{\left[p_1 + p_2 e^{(\lambda_1 - \lambda_2)t}\right]^2} .$$

$$h_T''(t) = p_1 p_2(\lambda_2 - \lambda_1)^3 e^{(\lambda_1 - \lambda_2)t} \left\{ \frac{p_1 - p_2 e^{(\lambda_1 - \lambda_2)t}}{\left[p_1 + p_2 e^{(\lambda_1 - \lambda_2)t}\right]^3} \right\} .$$

Using the definition of $h_T(t)$, the following theorem results.

Theorem 1. <u>For</u> $0 < p_1 < 1$, <u>the relationship</u> $\lambda_1 < h_T(0) < \lambda_2$ <u>holds</u>.

Proof. Note that

$$h_T(0) = p_1\lambda_1 + p_2\lambda_2 = (1 - p_2)\lambda_1 + p_2\lambda_2 = \lambda_1 + p_2(\lambda_2 - \lambda_1). \quad \#$$

The following theorem is a special case of the general result shown in Ref. 1.

Theorem 2. <u>The failure rate</u> $h_T(t)$ <u>is a decreasing function</u>.

Proof. From the expression for $h_T'(t)$ given above it is seen that $h_T'(t) < 0$ for every t. #

Reference 4 indicates the general result of which the following is a special case.

Theorem 3. $\lim_{t\to\infty} h_T(t) = \lambda_1$.

Proof. The result follows directly from the expression for $h_T(t)$ and the assumption made above that $\lambda_1 < \lambda_2$. #

Theorem 3 is an interesting result since it means that, no matter how great the probability p_2 of choosing an element at time $t = 0$ from the subgroup with failure rate λ_2 may have been, the failure rate $h_T(t)$ eventually becomes as close to λ_1 as desired. An intuitive explanation of this phenomenon is that as the element chosen at random from the population is observed to have an ever increasing lifetime without failure it becomes ever more likely that it is an element with a low failure rate. But the lowest failure rate it can have is λ_1 since there were only two possible failure rates in the population from which the element was taken, and λ_1 was the least of these two. Intuitively, it would seem that the longer the element functions the more likely should its failure rate be λ_1. The limiting process of Theorem 3 reflects this likelihood.

The range of $h_T(t)$ is bounded.

Theorem 4. $\lambda_1 < h_T(t) < \lambda_2$ <u>for every t</u>.

Proof. The proof follows from combining the results of Theorems 1, 2, and 3. #

Some general statements can be made about the shape of the failure rate curve.

Theorem 5. *For $p_1 < p_2$, the function $h_T(t)$ has an inflection point. Before the time at which this inflection point occurs the function is concave, and it is convex after that time.*

Proof. The function $h_T(t)$ has an inflection point when its second derivative vanishes.

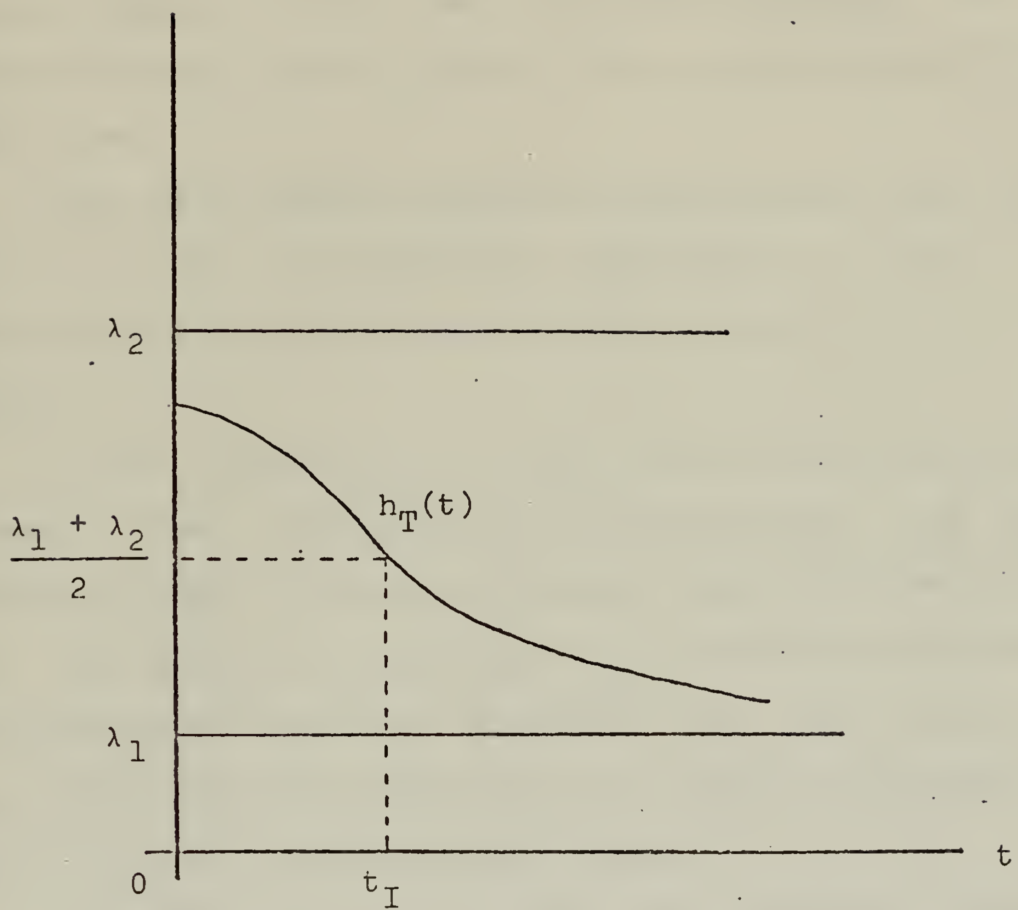
$$h_T''(t) = 0 \Leftrightarrow p_1 - p_2 e^{(\lambda_1 - \lambda_2)t} = 0 \Leftrightarrow t = \frac{1}{(\lambda_1 - \lambda_2)} \ln\left(\frac{p_1}{p_2}\right) .$$

For $p_1 < p_2$, this is a feasible time. That is, it is not a negative time.

Since $p_1 < p_2$, it is seen from the expression for $h_T''(t)$ that this function has a negative value at $t = 0$, and it is monotone increasing for $t > 0$. Therefore, $h_T(t)$ is concave before the time at which its inflection point occurs, and it is convex afterwards. #

The value of the failure rate function at the time the inflection point occurs is given by $h_T(t_I) = \frac{\lambda_1 + \lambda_2}{2}$, where $t_I = \frac{1}{(\lambda_1 - \lambda_2)} \ln\left(\frac{p_1}{p_2}\right)$.

As a result of Theorem 5 the graph of $h_T(t)$ for the general case of a mixture of two exponential failure rates has the form shown in Figure 2. An intuitive explanation for the shape of this curve can be given. Prior to the time at which the inflection point occurs $h_T(t)$ always has values closer to λ_2 than λ_1. It seems that if the element chosen at random has a relatively brief lifetime it probably has a high failure rate. The concavity of the curve, its shape tendency towards λ_2, prior to the inflection point seems to be an indicator that the element is more likely to have come from the subgroup with failure rate λ_2. The failure rate λ_2



Failure Rate Function Resulting From a Mixture of Two Exponential Types ($p_1 < p_2$)

Figure 2

seems to have a stronger attraction for the curve than λ_1 does prior to the inflection point.

At the inflection point, the failure rate function $h_T(t)$ has been shown to have a value exactly midway between λ_1 and λ_2. The intuitive explanation for the inflection point occurring when this happens is that the element is probably equally likely to have come from either subgroup. The failure rates λ_1 and λ_2 seem to have an equal attraction for the curve.

After the time the inflection point occurs, the curve becomes convex. The intuitive explanation for this is similar to that given above for the concave portion of the curve.

For the case when $p_2 < p_1$, $h_T''(t)$ is zero only at a negative time. And so for all feasible times, $t \geq 0$, the function $h_T(t)$ is convex and closer to λ_1 in value than to λ_2. The intuitive explanation for this involves the same reasoning used above for the case $p_1 < p_2$. The case $p_1 = p_2 = \frac{1}{2}$ is the same except that $h_T(t)$ has an inflection point at $t = 0$, and it then becomes convex.

A precise understanding about the conditional probability that an element randomly selected has come from a particular one of the two subgroups given that the element is still functioning at a certain time requires the definition of a residual mixing probability, called a "mixing probability" hereafter. Though the definition in this section is for the two exponentials case, it will in later sections have an

easy extension to more complex mixtures. By the mixing probability at time t is meant the function $P_i(t)$, i = 1 or 2, such that $P_i(t) = P(A_i | T > t)$, $t \geq 0$. This may be rewritten as

$$P_i(t) = P(A_i | T > t) = \frac{P(A_i, T > t)}{P(T > t)} = \frac{P(T > t | A_i)P(A_i)}{P(T > t)} = \frac{p_i e^{-\lambda_i t}}{\overline{F}_T(t)} .$$

For the particular case of i = 1, this means that

$$P_1(t) = \frac{p_1 e^{-\lambda_1 t}}{p_1 e^{-\lambda_1 t} + p_2 e^{-\lambda_2 t}} = \frac{p_1}{p_1 + p_2 e^{(\lambda_1 - \lambda_2)t}} .$$

Similarly,

$$P_2(t) = \frac{p_2}{p_2 + p_1 e^{(\lambda_2 - \lambda_1)t}} .$$

Note that $P_1(0) = p_1$ and $P_2(0) = p_2$.

Once the mixing probabilities $P_1(t)$ and $P_2(t)$ have been expressed in the above forms, it is easy to see that the intuitive explanation given above for the shape of the $h_T(t)$ curve is correct. Since $\lambda_1 < \lambda_2$, $P_1(t)$ is monotone increasing to the value 1, and $P_2(t)$ is monotone decreasing to a value of 0. And for the case where $h_T(t)$ has an inflection point (the case when $p_1 < p_2$), it can be seen that the following relationships hold:

$$P_1(t) < P_2(t) \text{ if } t < t_I$$

$$P_1(t) = P_2(t) \text{ if } t = t_I$$

$$P_1(t) > P_2(t) \text{ if } t > t_I .$$

The mixing probabilities can be used in another way to see how the curve $h_T(t)$ arises.

$$h_T(t) = \frac{p_1\lambda_1 e^{-\lambda_1 t} + p_2\lambda_2 e^{-\lambda_2 t}}{p_1 e^{-\lambda_1 t} + p_2 e^{-\lambda_2 t}} = \lambda_1 P_1(t) + \lambda_2 P_2(t) = E_t(\Lambda).$$

Here Λ is the random variable taking values λ_1 and λ_2 at time t with probabilities $P_1(t)$ and $P_2(t)$, respectively. So once these conditional probabilities are known, the function $h_T(t)$ is determined as an expected value. This is another fact having extension to cases more complex than a mixture of two exponentials, as will be seen in later sections.

The first derivative $h_T'(t)$ can be represented in terms of expected values at any time t too.

$$\frac{dP_i(t)}{dt} = \frac{\left(\sum_{j=1}^{2} p_j e^{-\lambda_j t}\right)(-\lambda_i p_i e^{-\lambda_i t}) + (p_i e^{-\lambda_i t})\left(\sum_{j=1}^{2} \lambda_j p_j e^{-\lambda_j t}\right)}{\left(\sum_{j=1}^{2} p_j e^{-\lambda_j t}\right)^2}$$

$$= \{E_t(\Lambda) - \lambda_i\} P_i(t) \text{ for } i = 1, 2.$$

And so

$$h_T'(t) = E_t'(\Lambda) = \lambda_1 P_1'(t) + \lambda_2 P_2'(t) = -\{E_t(\Lambda^2) - E_t^2(\Lambda)\}$$

$$= - \mathrm{Var}_t(\Lambda).$$

This is another way of showing that $h_T'(t) < 0$.

The second derivative $h_T''(t)$ also can be put in a concise expected value form.

$$h_T''(t) = -\{E_t'(\Lambda^2) - 2E_t(\Lambda)E_t'(\Lambda)\}$$

$$= -\{\sum_{i=1}^{2} \lambda_i^2 P_i'(t) - 2E_t(\Lambda) \sum_{i=1}^{2} \lambda_i P_i'(t)\}$$

$$= - \sum_{i=1}^{2} \{\lambda_i^2 - 2\lambda_i E_t(\Lambda)\} P_i'(t)$$

$$= \sum_{i=1}^{2} \{\lambda_i^2 - 2\lambda_i E_t(\Lambda)\}\{\lambda_i - E_t(\Lambda)\} P_i(t)$$

$$= E_t\{[\Lambda^2 - 2\Lambda E_t(\Lambda)][\Lambda - E_t(\Lambda)]\}$$

$$= E_t\{\Lambda^3 - 3\Lambda^2 E_t(\Lambda) + 2\Lambda E_t^2(\Lambda)\}$$

$$= E_t\{\Lambda - E_t(\Lambda)\}^3 .$$

The expression of $h_T''(t)$ as the third central moment of Λ will also have extension to general mixtures of exponentials.

III. MIXTURE OF THREE EXPONENTIAL TYPES (DISCRETE PARAMETERS)

Before considering in detail the general case of a mixture of n exponential types, some results which have been found for a population composed of three exponential types are presented. Function definitions for the n exponentials case are given first, and then in the remainder of this section it is assumed that n = 3. The more general definitions are needed in the next section. The definition of A_i, i = 1, 2, ..., n, parallels that given in the last section for A_1 and A_2. Also it will be assumed that the failure rate functions of the population subgroups satisfy $\lambda_1 < \lambda_2 < \ldots < \lambda_n$.

$$\bar{F}_T(t) = P(T > t) = \sum_{i=1}^{n} P(T > t | A_i)P(A_i) = \sum_{i=1}^{n} p_i e^{-\lambda_i t} .$$

$$f_T(t) = -\bar{F}_T'(t) = \sum_{i=1}^{n} p_i \lambda_i e^{-\lambda_i t} .$$

$$h_T(t) = \frac{\sum_{i=1}^{n} p_i \lambda_i e^{-\lambda_i t}}{\sum_{i=1}^{n} p_i e^{-\lambda_i t}} .$$

When the definition of the mixing probabilities $P_i(t)$ given in the last section is extended to populations where

the number of subgroups is $n > 2$, the failure rate function can once again be written as an expected value.

$$h_T(t) = \sum_{i=1}^{n} \lambda_i P_i(t) = E_t(\Lambda).$$

Using this fact and changing the upper index on all summation signs in the last section from 2 to n, it is apparent that in the general case

$$h_T'(t) = -\mathrm{Var}_t(\Lambda).$$

$$h_T''(t) = E_t\{\Lambda - E_t(\Lambda)\}^3.$$

A program was written in FORTRAN and run on the IBM 360 as a means of observing the behavior of $h_T(t)$ during various time intervals. For $n = 3$, examples checked with the aid of the program led to a conjecture which has been proven to be true.

Theorem 6. _If the population is composed of a mixture of three exponentials, with_ $\lambda_2 = \lambda_1 + a$, $\lambda_3 = \lambda_1 + 2a$ $(a > 0)$, _and_ $p_1 \leq p_3$, _then_ $h_T(t)$ _has an inflection point at_ $t_I = \frac{1}{2a} \ln\left(\frac{p_3}{p_1}\right)$. And $h_T(t_I) = \lambda_2$.

Proof. Note that at time t_I, which is feasible since $p_1 \leq p_3$,

$$P_1(t_I) = P_3(t_I).$$

Therefore,

$$\begin{aligned} h_T(t_I) = E_{t_I}(\Lambda) &= (\lambda_2 - a)P_1(t_I) + \lambda_2 P_2(t_I) + (\lambda_2 + a)P_3(t_I) \\ &= \lambda_2\{P_1(t_I) + P_2(t_I) + P_3(t_I)\} = \lambda_2 . \end{aligned}$$

And also,

$$h_T'''(t) = E_t\{\Lambda - E_t(\Lambda)\}^3$$

$$= \{(\lambda_2 - a) - \lambda_2\}^3 P_1(t) + (\lambda_2 - \lambda_2)^3 P_2(t)$$

$$+ \{(\lambda_2 + a) - \lambda_2\}^3 P_3(t)$$

$$= a^3\{P_3(t) - P_1(t)\} = 0. \qquad \#$$

Theorem 6 details the case for n = 3 which is analogous to the n = 2 case, where it was shown in the last section that the inflection point, if there is one, occurs at that time t when $h_T(t)$ is exactly midway between λ_1 and λ_2. Theorem 6 describes the case for n = 3 when the inflection point will occur exactly midway between λ_1 and λ_3. And at the time this inflection point occurs, for both the n = 3 and n = 2 cases, the least and greatest indexed mixing probabilities are equal. Even the forms for the two cases of the times at which this occurs are analogous. For n = 2 it was shown that, for $p_1 \leq p_2$, the inflection point occurs when

$$t = \frac{1}{(\lambda_2 - \lambda_1)} \ln\left(\frac{p_2}{p_1}\right).$$

Theorem 6 says the inflection point occurs when

$$t = \frac{1}{2a} \ln\left(\frac{p_3}{p_1}\right) = \frac{1}{(\lambda_3 - \lambda_1)} \ln\left(\frac{p_3}{p_1}\right).$$

It would be nice to find that $h_T(t)$ always has an inflection point when it takes on a value midway between λ_1 and λ_n. But this result does not carry over to more complex mixtures, and it fails to hold even for $n = 3$ when not all of the conditions of Theorem 6 are satisfied. As an example, consider the case where

$$\begin{array}{ll} \lambda_1 = 1 & p_1 = .2 \\ \lambda_2 = 2 & p_2 = .2 \\ \lambda_3 = 4 & p_3 = .6\ . \end{array}$$

Here 2.5 is midway between λ_1 and λ_3, but the inflection point was found to occur when $.24 < t < .25$ and $2.5739 < h_T(t) < 2.5915$.

Theorem 6 does not extend to the case $n = 4$, for instance, even when conditions analogous to those required in the theorem are satisfied. As an example,

$$\begin{array}{ll} \lambda_1 = 1 & p_1 = .1 \\ \lambda_2 = 2 & p_2 = .2 \\ \lambda_3 = 3 & p_3 = .3 \\ \lambda_4 = 4 & p_4 = .4\ . \end{array}$$

The inflection point was found to occur when $.4469 < t < .4470$ and $2.50955 < h_T(t) < 2.50960$.

A particular property applicable here of a general result proven in the next section is that $P_1(t)$ approaches the value 1 while $P_2(t)$ and $P_3(t)$ approach 0 as t becomes infinite. So

just as for the case $n = 2$, the least indexed mixing probability has 1 as a limit with increasing time, and all other mixing probabilities have 0 as a limit.

IV. GENERAL MIXTURE OF EXPONENTIAL TYPES (DISCRETE PARAMETERS)

The general case in which the original population consists of n subgroups, with all elements in any subgroup having the same exponential failure rate differing from the failure rate associated with every other subgroup, has some interesting properties. Definitions of p_i and $P_i(t)$ carry over directly to the general case from the n = 2 and n = 3 cases. In the last section expressions were given for $\overline{F}_T(t)$, $h_T(t)$, $h_T'(t)$, and $h_T''(t)$ for this mixture of n exponential types.

Theorem 3 is just a particular case of a property possessed by a general mixture of exponentials.

Theorem 7. $\lim_{t\to\infty} h_T(t) = \lambda_1$.

Proof.

$$h_T(t) = \frac{p_1\lambda_1 + \sum_{i=2}^{n} p_i\lambda_i e^{(\lambda_1 - \lambda_i)t}}{p_1 + \sum_{i=2}^{n} p_i e^{(\lambda_1 - \lambda_i)t}} . \qquad \#$$

Reference 1 shows that the function $h_T(t)$ is monotone decreasing. This of course means that $h_T'(t) < 0$ for all t.

When n = 2, $h_T(t)$ has at most one inflection point. This result has not yet been proven for the general case, even though computer runs of examples with various values of n have all revealed at most one inflection point for $h_T(t)$.

If it could be proven that $h_T''(t)$ is monotone increasing, then it would follow immediately that $h_T(t)$ has at most one inflection point. But this does not hold in general. Consider, for instance, an example where $n = 5$ and the parameter values for the mixture are

$$\begin{array}{ll} \lambda_1 = 2 & p_1 = .1 \\ \lambda_2 = 3 & p_2 = .1 \\ \lambda_3 = 4 & p_3 = .1 \\ \lambda_4 = 5 & p_4 = .3 \\ \lambda_5 = 6 & p_5 = .4\ . \end{array}$$

Here $h_T''(t)$ is negative at $t = 0$, increases to a positive maximum at $t = .824$, and then begins to decrease. All that can be concluded is that if $h_T(t)$ is concave at $t = 0$ it must have at least one inflection point because $h_T(t)$ approaches λ_1 as a limit as t becomes infinite.

A study of mixing probabilities in the general case reveals a number of interesting properties. Writing $P_i(t)$ in the form

$$P_i(t) = \frac{p_i}{\sum_{j=1}^{n} p_j e^{(\lambda_i - \lambda_j)t}} ,$$

it can be seen that

$$\lim_{t\to\infty} P_i(t) = \begin{cases} 1 \text{ for } i = 1 \\ 0 \text{ for } i = 2, \ldots, n. \end{cases}$$

This is an intuitively plausible result in light of the limiting behavior of $h_T(t)$ discussed above. The longer the element selected at time $t = 0$ continues to function, the closer its failure rate $h_T(t)$ gets to λ_1. It is reasonable, then, to expect that the conditional probability that the element selected came from the subgroup with failure rate λ_1, given continuing survival, should become greater with increasing time. And since the $h_T(t)$ curve, for large t at least, continually moves farther from the failure rates $\lambda_2, \ldots, \lambda_n$, it is reasonable that the conditional probabilities that the element selected has come from a subgroup with a failure rate greater than λ_1 should become less and less.

Additional facts about the behavior of the mixing probabilities can be determined.

Theorem 8. <u>If</u> $h_T(0) > \lambda_i$, $P_i(t)$ <u>reaches a maximum when</u> $h_T(t) = \lambda_i$.

Proof. An immediate extension to the general case of a result shown for n = 2 is

$$\frac{dP_i(t)}{dt} = \{h_T(t) - \lambda_i\}P_i(t).$$

This means that

$$\frac{dP_i(t)}{dt} > 0 \Longleftrightarrow h_T(t) > \lambda_i$$

$$\frac{dP_i(t)}{dt} = 0 \Longleftrightarrow h_T(t) = \lambda_i$$

$$\frac{dP_i(t)}{dt} < 0 \Leftrightarrow h_T(t) < \lambda_i \ . \qquad \#$$

The last line of the above proof shows that if $h_T(0) < \lambda_i$ the function $P_i(t)$ has its maximum at $t = 0$ and is monotone decreasing thereafter since $h_T(t)$ is monotone decreasing.

The intuitive explanation for Theorem 8 seems to follow from the monotone property of $h_T(t)$. Since $h_T(t)$ is always decreasing, it must, if $h_T(0) > \lambda_i$, get closer and closer to assuming the value λ_i until the time when $h_T(t) = \lambda_i$. It can be expected, then, that the conditional probability the element chosen came from the subgroup with failure rate λ_i, given survival until time t, should be increasing to a maximum at the point in time when $h_T(t) = \lambda_i$. As time continues beyond this, $h_T(t)$ gets farther and farther from λ_i. It would seem it must be getting less likely that the element chosen came from the subgroup whose elements have failure rate λ_i.

A property expressed in the proof of Theorem 6 for the case $n = 3$ does not carry over to an arbitrary exponential mix. In Theorem 6 it was shown that there is a case for $n = 3$ such that when $h_T(t)$ takes on a value midway between λ_1 and λ_n the total mixing probability mass for λ_i above this midvalue is equal to the total mixing probability mass below the midvalue. That is to say, when $h_T(t)$ reaches the midvalue, the element is equally likely to have come from a subgroup with failure rate greater than the midvalue as it

is to have come from one with failure rate less than the midvalue. But the example for $n = 5$ used earlier in this section shows this result does not carry over to the general case, even though the λ_i's form an arithmetic progression just as the λ_i's in Theorem 6 did. In the example, $h_T(t) = \lambda_3$ at $t = .38276$. But mixing probabilities at this time for λ_i values below the midvalue λ_3 satisfy $P_1(t) + P_2(t) = .42433$, while mixing probabilities for λ_i values above the midvalue satisfy $P_4(t) + P_5(t) = .45834$.

V. CONTINUOUS PARAMETER EXPONENTIAL MIXTURES

So far, the only populations which have been considered are those with a discrete number of subgroups and, thus, a discrete number of λ_i parameters. In this section it will again be assumed that each element in the population has exponential lifetime indexed by a failure rate parameter, but now the parameter will not be restricted to discrete values only. Instead, it will be assumed that it is a continuous, nonnegative random variable Λ with cumulative distribution function $G_\Lambda(\lambda)$ and density $g_\Lambda(\lambda)$. When $G_\Lambda(\lambda)$ is a step function, with $G_\Lambda(\lambda_n) = 1$ for finite λ_n, the result is one of the cases already considered.

Definitions in this continuous case are extensions of the ones in discrete cases. For instance,

$$\bar{F}_T(t) = P(T > t) = \int_a^\infty P(T > t \mid \Lambda = \lambda) g_\Lambda(\lambda) d\lambda = \int_a^\infty e^{-\lambda t} g_\Lambda(\lambda) d\lambda,$$

where $a > 0$ and $G_\Lambda(a) = 0$.

A number of the general properties of $h_T(t)$ are the same here as when the parameters were discrete. Reference 1 shows that $h_T(t)$ is monotone decreasing. Reference 4 indicates that $\lim_{t\to\infty} h_T(t)$ is the minimum value that the random variable Λ can assume. And at time $t = 0$, $h_T(t)$ is the expected value of the random variable Λ.

Theorem 9. $h_T(0) = E(\Lambda)$.

Proof. Note that

$$\bar{F}_T(t) = \int_a^{\infty} e^{-\lambda t} dG_\Lambda(\lambda) = E(e^{-\Lambda t}) .$$

$$f_T(t) = -\bar{F}_T'(t) = E(\Lambda e^{-\Lambda t}) .$$

$$h_T(t) = \frac{E(\Lambda e^{-\Lambda t})}{E(e^{-\Lambda t})} .$$

And therefore

$$h_T(0) = E(\Lambda) . \qquad \#$$

The properties of $h_T(t)$ are illustrated in the following example.

Example. Suppose that Λ is uniformly distributed over the interval $[a, b]$, $a > 0$. Then

$$g_\Lambda(\lambda) = \frac{1}{b - a} .$$

$$\bar{F}_T(t) = \int_a^b \frac{e^{-\lambda t} d\lambda}{b - a} = \frac{e^{-at} - e^{-bt}}{(b - a)t} .$$

$$f_T(t) = \frac{(1 + at)e^{-at} - (1 + bt)e^{-bt}}{(b - a)t^2} .$$

$$h_T(t) = \frac{(1 + at)e^{-at} - (1 + bt)e^{-bt}}{(e^{-at} - e^{-bt})t} .$$

Using elementary limiting arguments, it is found that

$$h_T(0) = \lim_{t\to 0} h_T(t) = \frac{a+b}{2} = E(\Lambda) .$$

And also

$$\lim_{t\to\infty} h_T(t) = a .$$

The analog in the continuous case of mixing probabilities is a conditional density function, $g_t(\lambda)$. This function gives the conditional density of the random variable Λ, given that the element selected at time $t = 0$ is still functioning at time t. For λ_i and λ_{i+1} such that $\Delta\lambda = \lambda_{i+1} - \lambda_i$ is small, $\{g_t(\lambda)\}\Delta\lambda$ will, then, give a good approximation to the probability that an element surviving until time t is one with failure rate Λ satisfying $\lambda_i \leq \Lambda \leq \lambda_{i+1}$. The conditional density is defined by

$$g_t(\lambda) = \frac{e^{-\lambda t} g_\Lambda(\lambda)}{\int_a^\infty e^{-ut} g_\Lambda(u)du} .$$

Note that, as in discrete cases, the failure rate function can be interpreted as an expected value at any time t since

$$h_T(t) = \frac{\int_a^\infty \lambda e^{-\lambda t} dG_\Lambda(\lambda)}{\int_a^\infty e^{-\lambda t} dG_\Lambda(\lambda)} = \int_a^\infty \lambda g_t(\lambda)d\lambda = E(\Lambda | T > t) = E_t(\Lambda).$$

Theorem 8, too, has its continuous analog.

Theorem 10. _If_ $h_T(0) > \lambda$, $g_t(\lambda)$ _reaches a maximum when_ $h_T(t) = \lambda$.

Proof. Note that

$$\frac{dg_t(\lambda)}{dt} = \frac{[\int_a^\infty e^{-ut}g(u)du][-\lambda e^{-\lambda t}g(\lambda)] + [e^{-\lambda t}g(\lambda)][\int_a^\infty ue^{-ut}g(u)du]}{[\int_a^\infty e^{-ut}g(u)du]^2}$$

$$= \{h_T(t) - \lambda\}g_t(\lambda) \quad . \qquad \#$$

And of course, as before, if $h_T(0) \leq \lambda$ the function $g_t(\lambda)$ has its maximum at $t = 0$ and is monotone decreasing thereafter since $h_T(t)$ is monotone decreasing.

VI. SOME NONEXPONENTIAL MIXTURES

In this section the requirement that all elements in the population have exponential failure rate is relaxed. Although a completely general mixture is not considered, previous theory is extended to mixtures more complex than the exponential ones considered thus far.

A useful function in studying more complex mixtures is the hazard function. This function has been used by others (see, for instance, Ref. 3). The hazard function of an element with random lifetime is the function $H(t)$ defined by $H(t) = -\ln \overline{F}(t)$, where $\overline{F}(t)$ is, as usual, the element's survival function. And if $f(t)$ and $h(t)$ are the element's density and failure rate functions, respectively, the following relationships hold.

$$h(t) = \frac{f(t)}{\overline{F}(t)} = H'(t) .$$

$$f(t) = -\overline{F}'(t) = H'(t)e^{-H(t)} .$$

And since

$$H(t) = \int_0^t h(s)ds ,$$

it follows that

$$f(t) = h(t)e^{-H(t)} .$$

Note that $H(t)$ is an increasing function since $\overline{F}(t)$ is decreasing.

Now suppose that the population is composed of n subgroups. All elements in subgroup i have density function $f_i(t)$, survival function $\bar{F}_i(t)$, failure rate function $h_i(t)$, and hazard function $H_i(t)$. None of these functions for a subgroup is equal to its corresponding function in any other subgroup. And again at time $t = 0$ each subgroup i makes up a proportion p_i of the total population. Without any other restrictions, this would describe the most general mixture of a discrete number of subgroups.

Consider now a case where each failure rate $h_i(t)$, $i > 1$, is greater than $h_1(t)$ by a constant at all times t. That is,

$$h_1(t) = h_1(t) + k_1$$

$$h_2(t) = h_1(t) + k_2$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$h_n(t) = h_1(t) + k_n ,$$

where $0 = k_1 < k_2 < \ldots < k_n$. Then

$$H_i(t) = \int_0^t h_i(s)ds = H_1(t) + k_i t .$$

And since $h_1(t) < h_2(t) < \ldots < h_n(t)$, $H_1(t) < H_2(t) < \ldots < H_n(t)$.

And also

$$f_i(t) = h_i(t)e^{-H_i(t)} = [h_1(t) + k_i]e^{-[H_1(t) + k_i t]} .$$

$$\bar{F}_i(t) = e^{-H_i(t)} = e^{-[H_1(t) + k_i t]} .$$

An element selected at random from the population at time t = 0 will, then, have random lifetime described by the failure rate function

$$h_T(t) = \frac{\sum_{i=1}^{n} p_i f_i(t)}{\sum_{i=1}^{n} p_i \bar{F}_i(t)} = \frac{\sum_{i=1}^{n} p_i [h_1(t) + k_i] e^{-[H_1(t) + k_i t]}}{\sum_{i=1}^{n} p_i e^{-[H_1(t) + k_i t]}} .$$

And since $F_i(0) = 0$ by assumption, it follows that

$$h_T(0) = \sum_{i=1}^{n} p_i h_i(0).$$

This corresponds to the result shown for exponential cases that $h_T(0)$ is the expected value, weighted by the p_i, of failure rates in the population at time t = 0. In fact, just as before, $h_T(t)$ can be interpreted at all times t as an expected value since now the mixing probabilities are given by

$$P_i(t) = \frac{p_i \bar{F}_i(t)}{\sum_{j=1}^{n} p_j \bar{F}_j(t)} = \frac{p_i e^{-[H_1(t) + k_i t]}}{\sum_{j=1}^{n} p_j e^{-[H_1(t) + k_j t]}} .$$

And so

$$h_T(t) = \sum_{i=1}^{n} h_i(t) P_i(t) .$$

$P_i(t)$ can actually be put into the same form that the mixing probability had in the exponential case since

$$P_i(t) = \frac{p_i e^{-k_i t}}{\sum_{j=1}^{n} p_j e^{-k_j t}} .$$

This means that the limiting result shown for exponential cases holds here. That is,

$$\lim_{t\to\infty} P_i(t) = \begin{cases} 1 \text{ for } i = 1 \\ 0 \text{ for } i = 2, \ldots, n. \end{cases}$$

Another limiting result holds too. If $h_T(t)$ is put in the form

$$h_T(t) = h_1(t) + \frac{\sum_{i=2}^{n} p_i k_i e^{-k_i t}}{\sum_{i=1}^{n} p_i e^{-k_i t}} ,$$

it can be seen that $\lim_{t\to\infty} [h_T(t) - h_1(t)] = 0$. Just as in the exponential case, $h_T(t)$ ultimately approaches as a limit the least failure rate function present in the original population.

Theorem 8, too, has its counterpart.

Theorem 11. <u>If</u> $h_T(0) > h_i(0)$, $P_i(t)$ <u>reaches a maximum when</u> $h_T(t) = h_i(t)$. <u>Otherwise</u>, $P_i(t)$ <u>is continually decreasing</u>.

Proof. The argument exactly parallels that used before, since

$$\frac{dP_i(t)}{dt} = \{h_T(t) - h_i(t)\}P_i(t) .$$

And $h_T(t)$ gets continually closer to $h_1(t)$ since

$$\frac{\sum_{i=2}^{n} p_i k_i e^{-k_i t}}{\sum_{i=1}^{n} p_i e^{-k_i t}} \quad \text{is a decreasing function.} \qquad \#$$

Consider now a population in which the failure rates for subgroups satisfy

$$h_1(t) = k_1 h_1(t)$$
$$h_2(t) = k_2 h_1(t)$$
$$\cdots\cdots\cdots\cdots\cdots$$
$$h_n(t) = k_n h_1(t) ,$$

where $1 = k_1 < k_2 < \ldots < k_n$. Note that in this case $\bar{F}_i(t) = [\bar{F}_1(t)]^{k_i}$.

The result in the exponential case that $h_T(t)$ is a decreasing function cannot be expected to carry over at all times here since $h_1(t)$ may at some times be increasing. However, $h_T(t)$ can be expressed as a time dependent multiple $y(t)$ of $h_1(t)$, and $y(t)$ is a decreasing function. To see this it is helpful to consider the case where the multipliers

are not discrete k_i's but instead come from a continuous random variable $K \geq 1$ with cumulative distribution function $G_K(k)$. When $G_K(k)$ is a step function, and $G_K(k_n) = 1$ for some finite k_n, the discrete case is described.

$$\bar{F}_T(t) = \int_1^\infty [\bar{F}_1(t)]^k dG_K(k) .$$

$$f_T(t) = \int_1^\infty k[\bar{F}_1(t)]^{k-1} f_1(t) dG_K(k) .$$

$h_T(t) = h_1(t)y(t)$, where the function $y(t)$ is

$$y(t) = \frac{\int_1^\infty k[\bar{F}_1(t)]^k dG_K(k)}{\int_1^\infty [\bar{F}_1(t)]^k dG_K(k)} = \frac{E[K\bar{F}_1(t)^K]}{E[\bar{F}_1(t)^K]} .$$

Note that

$$y'(t) \left\{\int_1^\infty [\bar{F}_1(t)]^k dG_K(k)\right\}^2$$

$$= \int_1^\infty [\bar{F}_1(t)]^k dG_K(k) \int_1^\infty k^2[\bar{F}_1(t)]^{k-1}[-f_1(t)] dG_K(k)$$

$$- \int_1^\infty k[\bar{F}_1(t)]^k dG_K(k) \int_1^\infty k[\bar{F}_1(t)]^{k-1}[-f_1(t)] dG_K(k).$$

This means that $y'(t) < 0$ if and only if

$$\left\{\int_1^\infty k[\bar{F}_1(t)]^k dG_K(k)\right\}^2 < \int_1^\infty [\bar{F}_1(t)]^k dG_K(k) \int_1^\infty k^2[\bar{F}_1(t)]^k dG_K(k).$$

But

$$\left\{\int_1^\infty k[\bar{F}_1(t)]^k dG_K(k)\right\}^2 = \left\{\int_1^\infty \left(\sqrt{[\bar{F}_1(t)]^k}\right)\left(k\sqrt{[\bar{F}_1(t)]^k}\right) dG_K(k)\right\}^2$$

$$< \int_1^\infty [\bar{F}_1(t)]^k dG_K(k) \int_1^\infty k^2[\bar{F}_1(t)]^k dG_K(k)$$

by the Cauchy-Schwarz inequality for Stieltjes integrals. This proves that y(t) is a decreasing function. And this means that once the $h_T(t)$ curve crosses below an $h_i(t)$ curve, it never intersects that $h_i(t)$ curve again. This result, too, was found in the exponentials mixture case.

Returning to the discrete mixture, $h_T(t)$ again approaches the least failure rate function present in the population as time increases.

Theorem 12.

$$\lim_{t\to\infty} \frac{h_T(t)}{h_1(t)} = 1 .$$

Proof. Using the definition of $h_T(t)$, it can be seen that

$$\lim_{t\to\infty} \frac{h_T(t)}{h_1(t)} = \lim_{t\to\infty} \frac{p_1k_1 + \sum_{i=2}^{n} p_ik_ie^{(k_1 - k_i)H_1(t)}}{p_1 + \sum_{i=2}^{n} p_ie^{(k_1 - k_i)H_1(t)}} = k_1 = 1 .$$

#

Noting that

$$P_i(t) = \frac{p_ie^{-k_iH_1(t)}}{\sum_{j=1}^{n} p_je^{-k_jH_1(t)}} ;$$

it can be seen that $h_T(t)$ has the usual expected value interpretation at any time t. That is,

$$h_T(t) = \sum_{i=1}^{n} h_i(t)P_i(t) = E_t[h_K(t)] .$$

Theorem 11 holds here also since once again

$$\frac{dP_i(t)}{dt} = \{h_T(t) - h_i(t)\}P_i(t) .$$

More research is needed to determine the existence of possible necessary and sufficient conditions which failure rate functions in the original population need to satisfy in order that properties of $h_T(t)$ and $P_i(t)$ seen for mixtures considered in this thesis hold.

BIBLIOGRAPHY

1. Barlow, R. E., Marshall, A. W., and Proschan, F., "Properties of Probability Distributions With Monotone Hazard Rate," The Annals of Mathematical Statistics, v. 34, p. 375-389, June 1963.

2. Proschan, F., "Theoretical Explanation of Observed Decreasing Failure Rate," Technometrics, v. 5, p. 375-383, August 1963.

3. Esary, J. D., Marshall, A. W., and Proschan, F., "Some Reliability Applications of the Hazard Transform," SIAM Journal of Applied Mathematics, v. 18, p. 849-860, June 1970.

4. Gnedenko, B. V., Belyayev, Y. K., and Solovyev, A. D., Mathematical Methods of Reliability Theory, Academic Press, 1969.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
3.	Assoc. Professor J. D. Esary, Code 55 Ey Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	1
4.	LT Michael Wayne O'Bar, USN 3220 Apache Avenue San Diego, California 92117	1
5.	Professor Sam C. Saunders Department of Mathematics Washington State University Pullman, Washington 99163	1
6.	Professor A. W. Marshall Department of Statistics University of Rochester Rochester, New York 14627	1
7.	Professor Z. W. Birnbaum Department of Mathematics University of Washington Seattle, Washington 98105	1
8.	Professor R. E. Barlow Department of Industrial Engineering and Operations Research University of California Berkeley, California 94720	1
9.	Professor Ernest M. Scheuer Management Science Department San Fernando State College Northridge, California 91320	1
10.	Professor Frank Proschan Department of Statistics The Florida State University Tallahassee, Florida 32306	1

11. Chief of Naval Personnel 1
Pers 11b
Department of the Navy
Washington, D. C. 20370

12. Naval Postgraduate School 1
Department of Operations Research
and Administrative Sciences
Monterey, California 93940

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

ORIGINATING ACTIVITY (Corporate author)

Naval Postgraduate School
Monterey, California 93940

2a. REPORT SECURITY CLASSIFICATION

Unclassified

2b. GROUP

REPORT TITLE

The Failure Rate and Residual Mixing Probabilities Resulting From a Mixture of Distributions

DESCRIPTIVE NOTES (Type of report and, inclusive dates)

Master's Thesis; September 1972

AUTHOR(S) (First name, middle initial, last name)

Michael Wayne O'Bar

REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
September 1972	44	4

8a. CONTRACT OR GRANT NO.

b. PROJECT NO.

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)

9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY

Naval Postgraduate School
Monterey, California 93940

13. ABSTRACT

Properties are developed for the failure rate function and residual mixing probabilities resulting from a mixture of distributions. After a detailed investigation of these properties for purely exponential mixtures, two more general mixtures are considered. Emphasis is placed on relating the current failure rate in the mixed population to the population's residual composition. The results have potential application in determining an appropriate period of "burn in."

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Reliability						
Failure Rate						
Mixture of Life Distributions						

DD FORM 1 NOV 65 1473 (BACK)

S/N 0101-807-6821

Thesis 135545
0165 O'Bar
c.1 The failure rate and residual mixing probabilities resulting from a mixture of distributions.